DOCUMENT RESUME

ED 065 172 PS 005 668

AUTHOR Stern, Carolyn
TITLE How Children Feel About Themselves: The Achilles Heel of Measurement.
INSTITUTION California Univ., Los Angeles. Early Childhood Research Center.
SPONS AGENCY Office of Economic Opportunity, Washington, D.C.
PUB DATE Dec 70
NOTE 14p.; Paper based on a presentation at the Annual meeting of the National Association for the Education of Young Children, Boston, Mass. (Nov. 1970)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Attitudes; Behavioral Science Research; Child Development; Classroom Observation Techniques; *Cognitive Development; *Early Childhood; Emotional Development; Evaluation Techniques; Intervention; *Measurement Instruments; Objectives; Preschool Children; Problem Solving; *Self Concept; Task Performance; *Testing

ABSTRACT

Two problems related to early childhood are studied: the specification of goals and the problem of measurement. Methods used to study these problems are to define objectives in the affective domain and to develop instruments to measure the attainment of these objectives. It is pointed out that the interrelationship between what the child is able to do and how he feels about himself is being more clearly recognized, and the lines of separation between the developmental and cognitive learning approaches are beginning to blur. It also pointed out that the cultivation of a positive self concept and the acquisition of cognitive skills must proceed in tandem and that both are important prerequisites for success in school and for the development of a competent, independent and contributing adult. There are two approaches to the assessment of behavior. The first method is observational, i.e., some scheme by which desired behaviors are categorized and rated by an external observer who is usually the classroom teacher or some other specially trained adult. The second technique relies on the subject's performance on specifically constructed tasks or test items. Finally, it is pointed out that there is a recognition of the need to find procedures for assessing change along both the emotional and cognitive dimensions so that the effectiveness of any preschool intervention can be more fully evaluated. (CK)

ED 065172

F-DEO
PS

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

HOW CHILDREN FEEL ABOUT THEMSELVES: THE ACHILLES HEEL OF MEASUREMENT[1]

by

Carolyn Stern

University of California, Los Angeles

December, 1970

PS 005668

[1]The work reported herein was carried out at the Early Childhood Research Center, Dr. Carolyn Stern, Director. This Center is funded by Grant CG 9938 from the Office of Economic Opportunity.

How Children Feel About Themselves: The Achilles Heel of Measurement[1]

Carolyn Stern, University of California, Los Angeles

Traditionally, the teacher of the young child has been concerned with social and emotional growth, whether the focus is on "socialization," i.e. developing the child's ability to relate to peers and adults in interpersonal contacts, or to "self-realization," i.e. actualizing his own inner potentialities and fulfilling himself as an individual. With the advent of intervention or compensatory preschool programs in the late fifties and early sixties, the effectiveness of child-centered nursery schools began to be seriously questioned and even attacked outright as being untenable for disadvantaged preschoolers.

The early studies which compared the effect of a planned curriculum with the "traditional" nursery usually demonstrated that any type of intervention was better than none, but that a structured, cognitively-oriented approach produced measurably superior gains in both I.Q. scores and academic competencies. The teacher for whom these were not the primary objectives of the early school experience insisted that her children were better adjusted and would easily catch up to and even surpass the performance of those who had been subjected to "robotizing" drill programs. Many studies presented the anomolous situation in which the goals of the intervention were expressed in terms of developing positive self-concept, but with the outcomes assessed in terms of tests which measured specific skill areas. Here the non-cognitive programs were at a disadvantage, but the teachers could legitimately claim that the tests didn't measure the right things.

---

[1]This paper is based on the presentation at the Annual Meeting of the National Association for the Education of Young Children, Boston, Massachusetts, November 1970.

There are here two separate problems. The first relates to the specification of goals, the second to that of measurement. With reference to the first problem, it is interesting to note the discrepancy between goals of preschool teachers, whether in the middle-class nursery or in compensatory programs such as Head Start, on the one hand, and those of kindergarten teachers and parents on the other. During the national evaluation, Head Start teachers were requested to rate a variety of goals in order of importance. These teachers consistently placed the highest value on affective, social-emotional adjustment, while giving the lowest ratings to acquisition of specific academic skills. This finding was corroborated in three separate but related research studies at the UCLA Head Start Evaluation and Research Center (Stern, 1970; Stern, Prichard & Rosenquist, 1970; and Stern, Kitano, Gaal, Goetz, & Ruble, 1970). These studies also indicated that parents of disadvantaged children, who are representative of the general public, still place the highest value on acquisition of school skills. It can be no surprise then that the tests which measure cognitive skills show disappointing increments over the program period, and that these findings are publicized as evidence that Head Start has failed. Unfortunately, changes in affective behaviors are difficult to assess because there are few instruments which measure these elusive behaviors with any acceptable degree of validity or reliability. A further complication is that the behaviors themselves have been described in vague, global terms; they must be translated into specific observable events which can be objectively evaluated.

The first step then is to define objectives[2] in the affective domain;

[2]At UCLA, a collaborative effort between the Early Childhood Research Center and the Center for the Evaluation of Instructional Programs is under way to develop a systematic taxonomy of goals for young children.

the second step would be to develop instruments to measure the attainment of these objectives. The question of measurement has become extremely critical with the current refocusing of interest on the feeling components of the child's development. The interrelationship between what the child is able to do and how he feels about himself is being more clearly recognized and the lines of separation between the developmental and cognitive learning approaches are beginning to blur. Even the most extreme proponents of the "developmental" approach would probably be willing to admit that a meaningful, positive self-concept cannot be developed without some real skills to back it up. Obviously, a child isn't going to feel very good about himself if he is constantly failing when he tries to do the things his peers are doing, or if he cannot meet the expectations of his parents and teachers. On the other hand, those who favor the cognitive emphasis would not deny that the child must feel comfortable with himself and have some confidence in his ability as a person so that he is willing to try new experiences and learn from them. There seems to be much more general agreement that the cultivation of positive self-concept and the acquisition of cognitive skills must proceed in tandem, and that both are important prerequisites for success in school and for the development of a competent, independent, and contributing adult.

Once the need for fostering both cognitive and affective competencies has been accepted, different types of programs need to be evaluated in terms of the degree to which they facilitate growth along these dimensions. While there are many useful instruments for measuring the acquisition of academic skills, the tools for assessment in the affective domain are extremely inadequate. The basic problem is to translate the multifaceted complex of affect into observable behaviors. With adults the self-report

technique has usually been found to be superior to other assessment devices. With children who have limited oral and written verbal skills, who have difficulty in following directions and in retaining statements in memory, the problems are immeasurably more complex. Yet some valid method of assessing behavior in this domain must be developed in order to implement and evaluate the effects of early schooling.

It is encouraging to note that many investigators are now directing their energies to the development of procedures for assessing the specific behaviors which comprise the affective domain. Even as these words are being written, new ways of assessing affect are being explored. Thus the survey of instruments, to which the rest of this paper is directed, is but a sampling of new developments in this area. For those interested in a more complete and exhaustive description of evaluation materials, two major resources are suggested: at the Educational Testing Service, Princeton, New Jersey, a new ERIC office has been established for the collection and dissemination of information about evaluation; at the University of California, Los Angeles, the Early Childhood Research Center and the Center for the Evaluation of Instructional Programs are collaborating in the writing of a book listing and evaluating published tests for young children.

There are basically two different approaches to the assessment of behavior. The first method is observational, i.e. some scheme by which desired behaviors are categorized and rated by an external observer who is usually the classroom teacher or some other specially-trained adult. The second technique relies on the subject's performance on specifically constructed tasks or test items.

Probably the most familiar type of measure in the first category is a teacher rating scale such as the Zigler Behavior Inventory, which was used in the 1966-67 National Head Start Evaluation. This measure consists of

50 items to which the teacher responds on a 4-point scale for each individual child. The items are grouped into nine sub-scales: sociability, independence, curiosity, persistence, emotionality, self-confidence, jealousy, achievement, and leadership. For the UCLA sample, which included a total of 148 children in 13 classes within five different delegate agencies, significant pre-post differences were found on only three of these scales; for independence and persistence there were reliable increases, but for achievement the significant difference in pre-post teacher ratings was in the opposite direction as had been anticipated. The most parsimonious explanation might be that most teachers were overly optimistic in the changes they expected to produce, so that their post-test ratings were depressed even though the achievement test scores showed that most children had actually made gains.

Because of the problem of the floating baseline, and other inadequacies of teacher rating scales, the Behavior Inventory was not used in the 1967-68 evaluation. Instead, the Social Interaction Observation (SIO), developed by Barbara Etzel and Russell Tyler at the University of Kansas, was employed. The SIO is a time-sampling procedure in which objectively observable interactions among children and adults are categorized and recorded. While this technique has many advantages, it is subject to two serious criticisms: first, it is not easy to administer since it requires highly skilled observers; and second, it is very difficult to interpret since norms are not available. Thus, even while the SIO was being used for the national evaluation, a task force of E & R Centers was assigned the responsibility for investigating, designing, and testing alternative procedures for assessing social-emotional behavior of preschool children.

As a result of this effort, two new instruments were used during the 1968-69 national evaluation. One of these was the Gumpgookies: A Test of

Motivation to Achieve, and the other a sociometric device called the Play Situation Picture Board. Both of these instruments fall into the second category, that is, procedures which evaluate performance on specifically-structured tasks.

The Gumpgookies was developed at the University of Hawaii Evaluation and Research Center by Bonnie Ballif and Dorothy Adkins. The title refers to a cartoon figure that looks something like Casper the Ghost. The child is told that he has his own little Gumpgookie who does whatever he does, and likes whatever he likes. Each item consists of a pair of these figures; the child is told a story in which one Gumpgookie's behavior is appropriate and motivated, and the other's behavior is disapproved and unmotivated. The child is told to select the Gumpgookie who is behaving most like the way he would behave in that situation.

When this measure was administered as a pretest, it proved to be too long (there were 100 items) and many of the statements were beyond the comprehension level of the beginning Head Start child. As a result, the children became restless and many of them either refused to complete the test or began to respond haphazardly, without even looking at the pictures. In response to these criticisms, 45 items were eliminated and the posttest consisted of only 55 items. Also, additional work has been done on the test by the University of Hawaii staff, and a revised 75-item test has been field tested and validated. However, the test is not available commercially and there is no funding for disseminating the test materials.

The Play Situation Picture Board was developed at the Michigan State University Evaluation and Research Center, under the direction of Robert Boger. This instrument requires that a polaroid picture of each child in the class, taken just prior to administering the test, be mounted on white fiberboard in four rows of four or five pictures. The child is first asked

to locate his own picture and identify himself with it. He is then asked to name all the other children. If he cannot produce each name himself, he is asked to point to the appropriate picture as the examiner gives the names of the children. After the child has demonstrated that he knows the other children in the class, he is shown a set of five pictures involving play situations and is asked to select three of them. When the child has made his selection, he is then asked to choose a child he would like to have engage in that activity with him.

There were several serious problems in the use of this instrument as a measure of change with a Head Start population. First of all, because of the high absentee rate, it was virtually impossible to obtain pictures of all the children within a reasonably limited period of time. Usually the test had to be administered before all members of the class had been photographed. A second serious problem was that the test provided that only children in the same class be included on the picture board. However, many day care or Head Start sites house two or more classes, sometimes in the same room, with a great deal of interaction among all the children. Thus the full range of options were not being offered.

The use of the instrument as a measure of change is also questionable. Obviously the child who is the most frequently selected on the pretest cannot improve his position and can only lose status; furthermore, the relative position of each child is highly dependent upon a stable classroom population. In Head Start, where many classes have 50% or more turn-over during the school year, the instrument is clearly inappropriate. Finally, retest reliability is very low, the test has not been validated, there are no norms, and the materials themselves are not commercially available.

There are a number of other instruments, many of them also not readily available, which attempt to assess affect in terms of the child's own

performance. One of the earliest of these is the Self Concept Referents Test, constructed by Bert Brown at the Institute of Developmental Studies, which is now affiliated with New York University.

Brown has based his test on the assumption that the child's perception of himself consists of two major components: self-as-object, i.e. how others such as parents, teachers, and peers view him, and self-as-subject, i.e. how he views himself. In the IDS test a polaroid picture is taken of the child and the child's identification with the picture established. He is then asked to describe himself by selecting one of a pair of bi-polar adjectives. There are 14 such pairs, including dirty-clean, good-bad, smart-dumb, etc. The test is repeated four times, with the child taking a different viewpoint each time (e.g. "Does the teacher think Johnny is good or bad?" "Do the other kids think Johnny is good or bad?" etc.).

This test was tried out in a pilot study at the UCLA E & R Center with 30 children from middle class homes; results showed a high clustering at the positive end of the scale, indicating that these children had already internalized socially-appropriate values. However, when the IDS was tested at the Michigan State University E & R Center with a Head Start population, approximately half of the children did not understand the meaning of the words used. Thus the instrument does not discriminate among advantaged preschool children and is inappropriate for use with disadvantaged preschoolers.

Rosastelle Wcolner has developed a Preschool Self-Concept Picture Test which consists of 10 pictures, each of which illustrates a pair of bipolar adjectives such as dirty-clean, active-passive, afraid-unafraid, etc. However, on this test the children are not required to know the words, only to identify with the picture representing the concepts. The main feature is that the child is also asked to select the picture of the child he would like

to be, after he has selected the one he feels he is most like. The test was validated by comparing the responses of normal children with those of children clinically identified as being disturbed. The level of correspondence between the perceived self and the ideal self clearly discriminated between these two groups.

While the author has carried out several studies with this instrument and is exceedingly optimistic as to its usefulness, field testing with a disadvantaged population found many discrepancies between the middle class norming group and the Head Start children. Further work would be necessary to determine whether the pictures failed to convey the same meaning, thus indicating lack of validity, or whether the test was discriminating a real difference in affect level in these two groups.

The Self-Social Constructs Test, developed by Henderson, Ziller, and Long, uses circles as symbols to represent the child, peers, and significant adults. The location of the circle the child selects to represent himself is presumed to be indicative of how the child views himself with respect to others. There are several types of items, e.g. esteem, social interest, identification, preference, etc. These item categories have been separately analyzed and in general seem to have good validity. The children have little trouble using the symbols in responding to the various test items. An adaptation of this method is being used by the Stanford Research Institute in its Follow Through evaluation.

Another instrument used in the SRI evaluation is adapted from the work of Crandall, Katkovsky, & Crandall, and is concerned with attribution of success or failure. Children are asked to respond to items such as, "I got a poor grade because the teacher doesn't like me." vs. "I got a poor grade because I didn't study." Evidently, children who have strong self-

esteem are able to take responsibility for their own performance, whereas those with weak self-concepts are not. The question of attribution has been extensively investigated by Weiner and his co-workers, with funding support from the UCLA Head Start Evaluation & Research Center.

Another test of school motivation is the one developed by Guy Strickland at the UCLA Center for the Study of Evaluation. This test is designed for children in the primary grades and is not appropriate for preschool or kindergarten. The items attempt to get at certain hypothesized attitudes towards school, and include attitudes toward specific subject matter, towards peers, and play activities. Contrary to expectation, analysis of the responses revealed only a generalized attitude which characterized all schoolwork, and a second factor relating to whether the exercise of authority in the classroom was structured-threatening or non-structured-accepting. This finding again points up the importance of the affective components in the learning environment.

Perhaps the most extensive set of measures for evaluation of the young child is the Cincinnati Autonomy Test Battery (CATB) developed by Thomas Banta. This test provides scores on nine separate subtests, and consists of variables which, it is hypothesized, make up the construct of autonomy. These include task initiation, curiosity, impulse control, intentional learning, innovative behavior, field independence, reflectivity, persistence, verbal and social competence, and resistance to distraction. The battery borrows liberally from the work of other investigators (Kagan, Maccoby, Witkin, et al.). While there are guidelines for scoring, no standardized norms are available. The CATB was considered by the Head Start social-emotional task force in their search for instruments in the affective doamin, but was not recommended because of poor

reliability. This battery (as well as a program developed to teach for autonomy) was used in a doctoral dissertation by Kay Kuzma and was found to have some limited usefulness.

At the UCLA Early Childhood Research Center several different lines of inquiry are proceeding in the attempt to establish valid and exportable techniques for assessing the child's feelings about himself. Similar research is being carried out throughout the country, where federally-funded studies, as well as the work of individual investigators, are directed at assessing this important aspect of human development. Many of these efforts are not yet ready for dissemination, but the technical literature is growing and new articles and books are beginning to appear.

The pendulum has swung from emphasis on social-emotional growth on a feeling level, as being something exceedingly important but not susceptible to measurement, to emphasis on cognitive growth, which can be more readily assessed. Now it has returned to an acceptance of the importance of the affective domain, but with this difference: there is a recognition of the need to find procedures for assessing change along both the emotional and cognitive dimensions so that the effectiveness of any preschool intervention can be more fully evaluated. Hopefully it will then be possible to utilize a wide variety of programs which demonstrate the desired progress in the child's total development.

REFERENCES

Adkins, D. & Ballif, B. Measurement of motivation to achieve in pre-school children. University of Hawaii Head Start Evaluation and Research Center, Honolulu, Hawaii. Final Report, 1968.

Banta, T. J. Tests for the evaluation of early childhood education: The Cincinnati Autonomy Test Battery. In J. Hellmuth (Ed.), Cognitive Studies, Volume 1, Seattle, Washington, 1970.

Boger, R. P. & Knight, S. S. Social-emotional task force. Michigan State University Head Start Evaluation and Research Center, East Lansing, Michigan. Final Report, 1969.

Brown, B. R. The assessment of self concept among four-year-old Negro and White children: a comparative study using the Brown-IDS Self Concept Referents Test. Institute for Developmental Studies, New York University, New York, New York, 1966.

Crandall, V. C., Katkovsky, W. & Crandall, V. J. Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations. Child Development, 1965, 36, 91-109.

Etzel, B. & Tyler, R. The social-interaction observation (S.I.O.) University of Kansas, Lawrence, Kansas, 1967.

Kukla, A. The cognitive determinants of achieving behavior. Doctoral dissertation, University of California, Los Angeles, 1970.

Kuzma, K. The effects of three preschool intervention programs on the development of autonomy in Mexican-American and Negro children. Doctoral dissertation, University of California, Los Angeles, 1970.

Long, B. H. & Henderson, E. H. Social schemata of school beginners: some demographic correlates. Merrill-Palmer Quarterly, 1970, 16, 305-324.

Long, B. H., Henderson, E. H. & Ziller, R. C. Developmental changes in the self-concept during middle childhood. Merrill-Palmer Quarterly, 1967, 13, 201-215.

Reed, L. Achievement motivation and self-attribution related to school achievement. Master's thesis, University of California, Los Angeles, 1970.

Stern, C. Teacher ratings of behavioral objectives as related to performance of children on specific tasks. Research report, University of California, Los Angeles, 1970.

Stern, C., Prichard, H. & Rosenquist, B. Teacher's expectations for achievement of children in Head Start (TEACH). Research report, University of California, Los Angeles, 1970.

Stern, C., Kitano, H., Gaal, A., Goetz, B., & Ruble, D. Application of group dynamics procedures to promote communication among parents and teachers. Research report, University of California, Los Angeles, 1970.

Strickland, G. Development of a school attitude questionnaire for young children. Center for the Study of Evaluation, University of California, Los Angeles, 1970.

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S. & Rosenbaum, R. Perceiving the causes of success and failure. In Jones, E. E., Kanouse, D., Kelley, H. H., Nisbett, R., Valins, S., and Weiner, B. (Eds.), Attribution: Perceiving the Causes of Behavior. In preparation (1971).

Woolner, R. Preschool self-concept picture test (PS-CPT). Unpublished manuscript, 1968.